

1

SYSTEM AND METHOD FOR DEPTH-RECOVERY OF A FACE IN AN IMAGE

RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Patent Applications Nos. 62/179,118, filed Apr. 28, 2015, 62/179,121, filed Apr. 28, 2015 and 62/179,132, filed Apr. 28, 2015, all of which is incorporated herein in their entireties.

GOVERNMENT INTEREST

This invention was made with government support under U.S. Department of Justice Award No. 2013-IJ-CX-K005. The U.S. government has certain rights in this invention.

BACKGROUND OF THE INVENTION

Automated recovery of 3D structural information from 2D images has received considerable attention from researchers over the last couple of decades. The first approaches to recovering 3D structure from 2D information include Shape-from-X techniques (Shading, Texture, Focus, etc). However, the complex appearance of skin under varying illumination makes face modeling using these approaches a very difficult task, and the results from these methods have largely been unsatisfactory. Multi-image techniques such as Structure from Motion or Stereo approaches utilize multiple views of the same face, separated either temporally or by pose, to reconstruct a 3D model, thereby limiting their real-world applicability. Photometric stereo based approaches have been demonstrated to obtain significantly accurate 3D face reconstructions, using multiple wavelengths of light for illumination and imaging. However, these techniques require a controlled illumination condition during acquisition.

Recently, shading information has been combined with generic shape information derived from a single reference model by utilizing global similarity of faces. However, this method is heavily dependent on the relevance of the template, requiring some form of manual initialization, and also the boundary conditions and parameters to be adjusted during the reconstruction process.

3D Morphable Models (3DMMs) are currently the most effective choice for reconstruction of 3D face models from a single image. The 3DMM technique is well understood as a powerful and reliable method for synthesis and analysis of 3D models of everyday objects such as faces. The formulation of the approach allows representation and rendering of a wide variety of 3D structures, textures, poses, and illuminations by controlling a few parameters. Perhaps the most impactful part of the technique is a method to automatically fit these parameters to an observed 2D rendering, hence allowing a complete and accurate reconstruction of 3D shape and texture from a single 2D image. The objective of the fitting procedure is formalized as the minimization of appearance dissimilarity computed in the rendered space, and it is solved by using an iterative stochastic gradient descent based method.

The 3DMM approach was revolutionary and unique when proposed. However, it suffers some draw-backs. To accurately fit a face, it requires manual initialization and oversight, and the iterative nature of the fitting technique makes it slow and therefore unusable for many applications requiring real-time performance. Additionally, the accuracy of the

2

3D reconstruction has never been thoroughly analyzed in literature; the only evaluation of the technique has been via the indirect route of evaluation of facial recognition across pose variations.

3DMMs demonstrated encouraging results from single input images, using separate linear shape and texture subspaces to describe the space of face models. While the technique is simple in formulation and impressive in reconstruction ability, it suffers from the requirement of manual initialization and the tediousness of the fitting procedure.

Recently, Generic Elastic Models (GEMs) were introduced as a new efficient method to generate 3D models from single 2D images. The underlying assumption in the GEM approach is that pure depth information is not significantly discriminative between individuals and it can be synthesized by using a deformable generic depth model, as long as the (x, y) spatial information of facial features is aligned. However, learning a generic 3D face model requires a large number of faces. Moreover, the use of loop subdivision to refine the mesh and densify the model results in an inhomogeneous distribution of vertices on the face, as shown in the middle face in FIG. 1.

Mesh refinement (densification) approaches are typically used in computer graphics and CAD tools to accomplish a similar goal. Previous notable attempts at computing dense correspondences between faces include optical-flow based techniques and adaptive meshing techniques. Perhaps the most popular technique for mesh refinement is loop subdivision, which has known uses for modeling faces, including 3D Generic Elastic Models (3D-GEM). Loop subdivision, and related mesh refinement techniques have two important negative aspects: (1) due to their formulation, they move the positions of the original fiducial points in an irrecoverable manner. This is a potential hazard that must be avoided for accurate resynthesis of the face from the representation. (2) These techniques are principally driven by subdividing the initial triangular mesh that is provided. In the case of faces, this initial triangular mesh is obtained from the fiducial points by means of Delaunay (or similar) triangulation technique, which results in numerous smaller triangles around dense fiducial point locations (such as the eyes and lips), and fewer, larger triangles around areas with sparser fiducial points, such as cheeks. The result is that after mesh refinement, the resulting mesh vertices are severely concentrated around certain areas of the face, leading to a non-homogeneous representation. An example of this is depicted in FIG. 1.

Therefore, it would be desirable to find a technique for densification that addresses the deficiencies of methods using loop subdivision for densification.

SUMMARY OF THE INVENTION

To address the problems associated with the loop subdivision method of densification, a novel refinement technique using interpolating splines is disclosed. This results in a set of points which is more homogeneous and which does not require the movement of the initial fiducial points.

Herein, the problem of depth-recovery is cast as a linear missing-data problem, given the observed structure and texture of the face. To do so, the face is represented using a dense 2D correspondence map, parameterized by a thin plate spline (TPS) framework. One of two linear data-completion techniques can be used to recover the missing depth information.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 shows a sample comparison of the vertex distribution of the result of two mesh refinement approaches. (L